Update of Model Probability for Release Rates of Radionuclides Using Bayes' Theorem

Jongtae Jeong*, Dong-Keun Cho, and Youn-Myoung Lee
Korea Atomic Energy Research Institute, 111, Daedeok-daero 989beon-gil, Yuseong-gu, Daejeon, Republic of Korea
*jtjeong@kaeri.re.kr

1. Introduction

Simulation models for a system are usually generated by conceptualizing the system and representing it in computer code. Model uncertainties are unavoidable during these processes. The quantification of model uncertainty based on model probability can be an efficient methodology for obtaining the degree of belief of a model. In this paper, we adopted the Bayes' theorem for the quantification of model uncertainty and applied it to simulation models for release rates of radionuclides from the radioactive waste repository.

2. Quantification of model uncertainty using Bayes' theorem

2.1 Bayes' Theorem

The way to update prior probability of model M_k into posterior probability using Bayes' theorem for a set of models and experimental data D is given by the following equation [1]:

$$\Pr(M_k|D) = \frac{\Pr(M_k)\times Pr(D|M_k)}{\sum_{i=1}^{K}\Pr(M_i)\times Pr(D|M_i)}\ , k = 1, \ldots, K \quad (1)$$

Where, $\Pr(M_k)$ is prior probability of model M_k, $\Pr(M_k|D)$ is posterior probability of model M_k, and $\Pr(D|M_k)$ represents likelihood of model M_k given observed data D.

A common formulation for a model prediction can be written as follows:

$$y = f_k + \varepsilon_k\ , \varepsilon_k \sim N(0, \sigma_k^2) \quad (2)$$

Where y is a system response, and f_k is the prediction of y by a model M_k. ε_k is the error for both bias associated with model prediction f_k of response y and measurement error. ε_k is assumed to be an independent and identically distributed normal variable with zero mean and a constant variance. σ_k is the standard deviation of the error.

Eq.(2) can be represented in a probability distribution form as

$$g_Y(y|M_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} exp\left(-\frac{(y-f_k)^2}{2\sigma_k^2}\right) \quad (3)$$

$g_Y(y|M_k)$ is the predictive distribution of response y under model M_k. Using the above equation, the likelihood function of σ_k for each model M_k given a data set d_n is expressed by

$$\Pr(d_n|M_k, \sigma_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(d_n - f_{kn})^2}{2\sigma_k^2}\right) \quad (4)$$

Where, f_{kn} is the prediction of data d_n by model M_k.

Because experimental data are independent of one another, the likelihood of σ_k for each model M_k can be calculated by multiplying $\Pr(d_n|M_k, \sigma_k)$ in the above equation as represented by

$$\Pr(D|M_k, \sigma_k) = \left(\frac{1}{2\pi\sigma_k^2}\right)^{N/2} exp\left(-\frac{\sum_{n=1}^{N}(d_n - f_{kn})^2}{2\sigma_k^2}\right) \quad (5)$$

Model likelihood $\Pr(D|M_k)$ is expressed by marginal likelihood integral as follows:

$$\Pr(D|M_k) = \int \Pr(D|M_k, \sigma_k)\, g(\sigma_k|M_k) d\sigma_k \quad (6)$$

In general, the maximum likelihood estimation is implemented to evaluate model likelihood $\Pr(D|M_k)$ instead of finding a direct solution of Eq.(6). That is, taking the derivative of the logarithm of Eq. (5) with respect to σ_k and setting it equal to zero, and solving the equation for σ_k gives

$$\sigma_k^2 = \frac{\sum_{n=1}^{N} \varepsilon_{kn}^2}{N}, \varepsilon_{kn} = d_n - f_{kn} \quad (7)$$

By putting the above equation into the exponential term in Eq.(5), likelihood $Pr(D|M_k)$ of each model M_k given a set of experimental data D is computed:

$$\Pr(D|M_k) = \left(\frac{1}{2\pi\sigma_k^2}\right)^{N/2} exp\left(-\frac{N}{2}\right) \quad (8)$$

2.2 Application to Release Rates of Radionuclide

KAERI developed a GS-TSPA code for the post-closure safety assessment of a radioactive waste repository [2]. In this code, three models for the release rate of radionuclides from the waste canister are considered; annual release rate, congruent release, and surface release. The results of release rate for I-129 by three models and hypothetical experimental data are summarized in Table 1. The hypothetical experimental data are assumed to be the mean value of simulated results using three models because experimental data are not available.

Table 1. Release rates of I-129 (g/yr)

Time (yr)	Exp. Data	Annual Release	Congruent Release	Surface Release
10,000	1.60E-01	2.97E-02	2.96E-02	4.22E-01
20,000	7.65E-02	5.37E-03	4.38E-03	2.20E-01
30,000	3.02E-03	1.20E-03	1.78E-04	7.68E-03
40,000	4.50E-04	1.06E-03	2.44E-05	2.69E-04
50,000	3.59E-04	1.05E-03	1.61E-05	9.39E-06
60,000	3.55E-04	1.05E-03	1.42E-05	3.28E-07
70,000	3.53E-04	1.05E-03	1.31E-05	1.15E-08
80,000	3.53E-04	1.05E-03	1.22E-05	4.01E-10
90,000	3.52E-04	1.04E-03	1.15E-05	1.40E-11
100,000	3.51E-04	1.04E-03	1.10E-05	4.90E-13

We assumed that prior probabilities for three models are uniformly distributed. The updated posterior probabilities are summarized in Table 2. As shown in Table 2, the annual release rate model is most likely to give closest predictions among the models considered. The significantly low posterior probability of surface release model indicates that the model fits the data very poorly.

Table 2. Prior and posterior probabilities for three models

	Annual Release	Congruent Release	Surface Release
Prior probability	1/3	1/3	1/3
Standard deviation	4.70E-2	4.72E-2	9.43E-2
Likelihood	1.30E+7	1.25E+7	1.25E+4
Posterior probability	0.51	0.49	4.88E-4

3. Conclusions

We adopted the Bayes' theorem for the quantification of model uncertainty based on model probability. According to the results of application example, we found that the Bayes' theorem can be used as an efficient tool for the quantification of model uncertainty. However, the real experimental data are necessary to obtain more exact degree of belief for models by applying Bayes' theorem.

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea Nuclear Research and Development Program (2017M2A8A5014856) funded by the Ministry of Science and ICT.

REFERENCES

[1] I. Park and R.V. Grandhi, "A Bayesian statistical method for quantifying model form uncertainty", Reliability Engineering and System Safety 129, 46-56 (2014).

[2] Y.M. Lee and Y.S. Hwang, "A Goldsim model for the safety assessment of a HLW repository", Progress in Nuclear Energy, 51 (2009).